a folha

Boletim da língua portuguesa nas instituições europeias

http://ec.europa.eu/translation/portuguese/magazine

N.º 59 — primavera de 2019

Partilhar recursos linguísticos no portal eTradução — repositório nacional de recursos de tradução

Equipa eTradução
Agência para a Modernização Administrativa (AMA)

O que é o serviço eTraducao.gov.pt?

O eTradução[(1)] é um serviço que resulta do projeto ELRI[(2)] (European Language Resource Infrastructure) cofinanciado pela União Europeia, que permite às instituições da Administração Pública portuguesa recolher e partilhar recursos linguísticos através de uma plataforma *online*.

O que são recursos linguísticos?

Recursos linguísticos são quaisquer textos escritos em formato digital (isto é, acessíveis usando um computador). Podem ser bilingues (por exemplo, documentos em português com o seu equivalente em inglês) ou monolingues (por exemplo, textos apenas em português).

Porquê partilhar recursos linguísticos?

Ao partilhar os seus recursos linguísticos, está a contribuir para a melhoria dos serviços públicos de tradução e também para garantir que mais serviços públicos nacionais e europeus estejam disponíveis em português. O resultado é uma maior presença da língua portuguesa na vida pública e uma maior «pegada digital» da nossa língua.

Além disso, estará a apoiar o desenvolvimento do eTranslation[(3)] — um sistema de tradução por máquina, que já está disponível gratuitamente, para uso na Administração Pública em toda a UE. Quanto mais recursos linguísticos de qualidade forem recebidos, mais eficaz se torna a tradução. Portanto, ao contribuir regularmente com recursos portugueses para o eTradução, estará a ajudar e a garantir que as traduções do eTranslation de/para português fiquem cada vez melhores e mais precisas.

Ao contrário dos serviços de tradução *online* gratuitos, o eTranslation é regularmente «treinado» com dados atualizados da Administração Pública. Isto significa que ao incluir os seus dados no treino passa a ser mais eficaz na tradução de textos que são importantes para si.

Que tipos de textos são os mais adequados para partilhar?

Exemplos dos tipos de documentos que pode partilhar:

- comunicados de imprensa;
- relatórios anuais;
- regulamentos, termos de serviço, políticas e procedimentos de recursos humanos, etc.;
- conteúdo de um *site*;
- guias, folhetos, panfletos, formulários;
- quaisquer outros documentos para os quais exista uma versão portuguesa.

Quais os formatos de ficheiro aceites para fazer* upload *no eTradução?

O eTradução está preparado para suportar vários tipos de ficheiros, nomeadamente:

- Microsoft Word e ficheiros de texto OpenDocument (.doc, .docx, .odt, .rtf);
- Microsoft Excel (.xls, .xlsx);
- PDF (.pdf);
- ficheiros de texto simples (.txt, .xml, .tbx);
- ficheiros de memória de tradução (.tmx, .sdltm);
- todos os ficheiros anteriores comprimidos em formato ZIP (.zip).

Quem pode aceder aos dados que partilhou?

O contribuidor tem o poder de decidir quem acede aos seus dados, existindo quatro grupos-padrão com os quais pode partilhá-los:

- a sua organização (todos os utilizadores da sua organização);
- organizações nacionais (organizações da AP portuguesa e universidades);
- organizações nacionais + Comissão Europeia (organizações nacionais e Comissão Europeia);
- o público (dados distribuídos pelo portal de dados abertos nacional Dados.gov[(4)]).

Se existir necessidade de mais níveis de partilha para recursos específicos, pode entrar em contacto com o suporte do eTradução[(5)].

Como preparar os seus dados?

Crie uma pasta partilhada para este propósito no seu computador ou na rede da sua instituição. De forma individual ou em conjunto com os seus colegas tradutores, pode periodicamente guardar documentos traduzidos. Dessa forma, quando quiser fazer o *upload* de recursos para o eTradução, os ficheiros estarão à mão e pode facilmente carregar uma série de ficheiros de uma só vez.

Certifique-se de que a versão em português é guardada num documento e a versão na outra língua é guardada num documento separado.

Certifique-se de que os textos que partilha não contêm informações pessoais que possam levar a que um indivíduo seja identificado ou identificável. Em caso de dúvida, entre em contacto connosco.

Não há restrições de tempo para partilhar — pode partilhar sempre que lhe for conveniente a si ou aos seus colegas. O que significa que é fácil reservar alguns minutos uma vez por semana ou uma vez por mês — conforme lhe for mais conveniente — para aceder e partilhar os seus recursos.

Os recursos carregados no eTradução, uma vez publicados, estarão disponíveis para *download* como um ficheiro de memória de tradução (.tmx) — assim, se trabalha com ferramentas de apoio à tradução (por exemplo, SDL Trados, MemoQ, Wordfast, OmegaT, etc.), pode melhorar o desempenho dessas ferramentas com estes recursos.

Lembre-se:

Com cada documento que enviar, está a apoiar a presença da língua portuguesa no panorama europeu, impedindo o seu declínio digital e melhorando a qualidade dos serviços de tradução para si e para todos os que trabalham na Administração Pública portuguesa. Está também a apoiar os Estados-Membros da UE a trabalharem em parceria e a colaborarem nos nossos vários idiomas.

Sabia que...

Pode fazer *upload* de conjuntos de ficheiros para o eTRADUÇÃO, até 100 MB de uma só vez?

Isso faz com que o processo de *upload* dos seus recursos seja rápido e fácil.

Obrigado pela sua contribuição inestimável para o sucesso do projeto eTradução!

etraducao@ama.pt

---

(1) *Repositório Nacional de Recursos de Tradução: eTradução*, https://etraducao.gov.pt/pt-pt/.
(2) *European Language Resource Infrastructure*, http://www.elri-project.eu/.
(3) Comissão Europeia, *eTranslation*, https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation.
(4) Agência para a Modernização Administrativa, *Portal de Dados Abertos da Administração Pública: dados.gov*, https://dados.gov.pt/pt/.
(5) Correio eletrónico eTradução: etraducao@ama.pt.

Anãos guardiães

Jorge Madeira Mendes
Antigo funcionário da Direção-Geral da Tradução — Comissão Europeia

Uma característica notória do português, suscetível de causar dores de cabeça, não só aos que o têm como língua-mãe, mas ainda mais a quem procura aprendê-lo, são os plurais dos substantivos terminados em *ão*.

Com efeito, temos m*ãos* e cidad*ãos* — mas p*ães*, c*ães*, burl*ões* e comich*ões*.

A explicação está nas formas arcaicas com terminações em *ano*, em *ane* e em *one*. O arcaísmo pode remontar ao latim ou a fases precoces do já então português, mas o que importa é que, logicamente, cada uma destas terminações formava o plural acrescentando um simples *s*, ou seja, em *anos*, *anes* e *ones*.

Ora, as três formas do singular convergiram — todas — para o atual *ão*, mas a evolução dos respetivos plurais manteve a diferença. E, assim, *ano/anos* evoluiu para *ão/ãos*, *ane/anes* evoluiu para *ão/ães* e *one/ones* evoluiu para *ão/ões*.
Por conseguinte, saber como se forma o plural de uma palavra cujo singular termina em *ão* torna-se fácil... se se conhecer a forma arcaica desse singular.

Para quem domine razoavelmente o castelhano — a língua mais próxima da nossa[(1)] —, uma via facilitadora consiste em consultá-lo. Nesta língua, com efeito, as terminações singulares *ano*, *an* (ou *án*) e *ón* mantiveram-se, não convergindo para nenhuma forma única, pelo que os respetivos plurais requerem apenas a adição de um *s* (ou de *es* se a terminação do singular for em *n*, e suprimindo o acento gráfico nas palavras cujo singular o comporte): *anos*, *anes* e *ones* (*ciudadano/ciudadanos*, *pan/panes*, *afán/afanes*, *pasión/pasiones*).

Esta mnemónica já me ajudou, por exemplo, a determinar o plural de *ancião*: lembrando-me que, em castelhano, o equivalente é *anciano*, bastou-me acrescentar um *s* à forma do singular português[(2)]. Entretanto, para meu desgosto, verifiquei que a edição de 1984 do *Dicionário da Língua Portuguesa* de Fernando J. da Silva[(3)] já avalizava a forma *anciões*, em paralelo com a forma, mais lógica, *anciãos*.

A maior complicação do português está a gerar um fenómeno de recurso acrítico e uniforme a *ões* quando não é evidente o modo de formar o plural de um substantivo cujo singular termine em *ão*.

Dois exemplos flagrantes são os plurais de *anão* e de *guardião*. Os *anões* entraram há muito e, pouco a pouco, os *guardiões* vão-se insinuando. Ora, as formas corretas destes plurais são, respetivamente, ***anãos*** e ***guardiães***. O erro *anões* está tão difundido que muitos dicionários optaram pela sua avalização, como alternativa válida a *anãos*. Já na minha infância (vão seis décadas) havia a história da «Branca de Neve e os sete *anões*»... a tal ponto que me custa assimilar a versão «Branca de Neve e os sete *anãos*» (que seria a correta).

Quanto aos *guardiões*, creio não terem ainda conquistado a aprovação oficial que o uso costuma granjear, mas, com base no que ouço e leio nos órgãos de comunicação portugueses, para lá caminhamos aceleradamente.

Note-se que, em espanhol, se diz *enano/enanos* e *guardián/guardianes*.

jorge.mendes909@gmail.com

À parte:

Não diga *rei-te-rar*. Diga *re-i-te-rar*.

Explicação:
A génese da palavra *reiterar* consiste na adição do prefixo *re* ao verbo *iterar*, ou seja, «reiterar» significará «iterar de novo».

Génese idêntica à da palavra *reiniciar* (de *re+iniciar*, que ninguém pronuncia *rei-ni-ci-ar*, mas sim *re-i-ni-ci-ar*), à da palavra *reidratar* (de *re+hidratar*, que ninguém pronuncia *rei-dra-tar*, mas sim *re-i-dra-tar*) ou à da palavra *reinventar* (de *re+inventar*, que ninguém pronuncia *rãe-ven-tar*, mas sim *re-in-ven-tar*).

A associação das vogais *e* e *i* em ditongo colide com a lógica presente na formação de *reiterar*. Esta palavra tem a mesma génese de *renovar*, *repor*, *refazer*, *repisar*, *redizer*, *reaver*, *recuperar*.

(1) Estou aqui a considerar que o galego e o português são, não tanto duas línguas distintas (exceto por conveniência política), mas antes variantes de uma mesma matriz idiomática.
(2) De assinalar que o nome da vila transmontana de Carrazeda de Ansiães nada tem que ver com pessoas idosas.
(3) Silva, F. J. da, *Dicionário da Língua Portuguesa*, Editorial Domingos Barreira, Porto, 1984.

Duxambé, Chechénia e os estados Xã e Chim

Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

> XII — O emprêgo de ch ou de x, os quais histórica e ainda dialectalmente não eram nem são idênticos no valor fonético, regula-se pela sua origem, e a consulta ao VOCABULÁRIO torna-se necessária. Deve ter-se em atenção que *ch* corresponde a cl, fl, pl, t'l [*sic*] latinos, e a *ch* francês nas palavras desta origem; *x* corresponde a *x* e a *s* latinos. Nos vocábulos de origem arábica o emprêgo de *x*, e não de *ch*, é de rigor; assim, *xeque*, e não *che(i)k.*
>
> *Formulário Ortográfico* de 1911[(1)]

Ch e x — relíquia fonética e ortográfica nos aportuguesamentos

O Formulário Ortográfico de 1911, junto com o respetivo Prontuário Ortográfico[(2)], foi o documento fundador da moderna ortografia portuguesa. Observando-se o parágrafo acima, extraído do formulário, verifica-se que a opção entre «**x**» e «**ch**» não deverá ser aleatória ou determinada por simples critérios estéticos, mas sim determinada pela **fonética** e **origem** das palavras.

Por que razão, nos **aportuguesamentos**[(3)], se escreve Xangai e não Changai, China e não Xina, xaile e não chaile, chador e não xador, champô e não xampô, Caxemira e não Cachemira?

Na generalidade das geografias do português, o «x» e o «ch» aparentam ser utilizados indistintamente para representar a fricativa pré-palatal surda, o fonema representado pelo símbolo /**ʃ**/ do alfabeto fonético internacional (AFI). Porém, em muitos dos vernáculos setentrionais (Minho, Trás-os-Montes e Alto Douro, Beira Interior e Litoral) mantém-se ainda o som /**tʃ**/ (africada pós-alveolar surda) — representado pelo dígrafo «ch» — ao lado do som /ʃ/ — representado pela letra «x». Esta diferença permanece clara também no galego atual e respetiva ortografia[(4)]. Confrontar, por exemplo, a pronúncia de «cheque» e «xeque» no *Vocabulário Ortográfico Português*[(5)] e as diferenças de ortografia e pronúncia no *Dicionário da Real Academia Galega*[(6)] e *Dicionario de pronuncia da lingua galega*[(7)]:

	pronúncia	**cheque**	**xeque**	**bucho**	**buxo**	**chá**	**xá**
pt	Díli	ʃˈɛ.kɨ	ʃˈɛ.kɨ	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	Lisboa (não padrão)	ʃˈɛk	ʃˈɛk	bˈu.ʃu	bˈu.ʃu	ʃˈa	ʃˈa
	Lisboa (padrão)	ʃˈɛ.kɨ	ʃˈɛ.kɨ	bˈu.ʃu	bˈu.ʃu	ʃˈa	ʃˈa
	Luanda	ʃˈɛ.kɨ	ʃˈɛ.kɨ	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	Maputo (não padrão)	ʃˈɛkʰ	ʃˈɛkʰ	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	Maputo (padrão)	ʃˈɛ.kɨ	ʃˈɛ.kɨ	bˈu.ʃu	bˈu.ʃu	ʃˈa	ʃˈa
	Rio de Janeiro (não padrão)	ʃˈɛ.kɪ	ʃˈɛ.kɪ	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	Rio de Janeiro (padrão)	ʃˈɛ.kɪ	ʃˈɛ.kɪ	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	São Paulo (não padrão)	ʃˈɛ.ki	ʃˈɛ.ki	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
	São Paulo (padrão)	ʃˈɛ.ki	ʃˈɛ.ki	bˈu.ʃʊ	bˈu.ʃʊ	ʃˈa	ʃˈa
gl	estándar	ˈt͡ʃɛke̞	ˈʃɛke̞	ˈbut͡ʃo̞	ˈbuʃo̞	— (té)	ˈʃa (xa)

Também o crioulo cabo-verdiano, baseado num português mais antigo, mantém essa diferença entre /ʃ/ e /tʃ/, embora a solução ortográfica adotada seja diferente das do português e do galego, pois prescinde totalmente do «ch» e adota «x» e «tx» (como o faz o catalão).

- xikra (xícara)
- txuba (chuva)

Se se recuar no português mais de trezentos anos, ter-se-á grafado Xangai, refletindo o som /ʃ/ no chinês, pronunciando-se [ʃɐ̃gˈaj]. A pronúncia e a ortografia mantiveram-se no português (tal como no galego), sendo a marca dessa pronúncia dada pelo «x». Já China (do persa چین) pronunciar-se-ia antigamente [tʃˈi.nɐ], pelo que se terá grafado China. Com o tempo, a pronúncia predominante passou a [ʃˈi.nɐ], mantendo-se a ortografia (o galego manteve tanto a pronúncia como a ortografia). Hancheu e Sucheu são outros aportuguesamentos tradicionais de cidades do litoral chinês, refletindo o «ch» o som /tʃ/ no chinês.

Nestas e noutras palavras mais antigas incorporadas no português, o aportuguesamento terá sido **feito e validado** por contacto direto dos portugueses primeiro com os árabes e posteriormente com vários outros povos, sobretudo ribeirinhos, dos continentes americano, africano e asiático. Havendo a possibilidade de registar os sons /ʃ/ e /tʃ/ existentes (então) no português e nas diferentes línguas de origem das palavras, é natural que os portugueses o tivessem feito usando os recursos próprios da sua ortografia(8). Ou seja, a solução tradicional adotada nos aportuguesamentos para distinguir os sons /ʃ/ e (o antigo) /tʃ/ foi:

- x — ʃ
- ch — tʃ

É curioso verificar que, enquanto o som /tʃ/ desaparece substituído por /ʃ/ no português europeu padrão, em certas variantes do português do Brasil o som /tʃ/ (re)aparece em sequências como «te» ou «ti» e que, em quase todas as variantes do português, o «s» e «z» nos finais de sílaba se transformou em /ʃ/(9). Por sua vez, no galego o «x» (/ʃ/) expandiu-se substituindo o som /ʒ/, representado em português pelo «j» ou «g» (antes de «e» ou «i»)(10).

No português de Portugal, o som /tʃ/, que desde há três séculos estava conotado com vernáculos do norte mais interior do país, volta a ser pronunciado por certas pessoas que pontuam frequentemente o seu discurso com termos ingleses (geralmente com pronúncia americana). Esta prática entra em conflito com a tendência, mais popular, de aportuguesar a pronúncia desses mesmos termos ingleses. Na *Infopédia*(11) podem consultar-se alguns exemplos dessas duas pronúncias nos dicionários de inglês-português e de língua portuguesa, respetivamente:

	en	**pt**
check-in	[ˈtʃekɪn]	[ʃɛˈkin]
chip (chipe)	[tʃɪp]	[ˈʃip]; [ˈʃip(ə)]

Ch e x nos aportuguesamentos mais recentes

> (…) x é tradicionalmente a letra que representa a consoante fricativa pré-palatal surda [ʃ], e não ch, que se reserva para a transliteração de grafemas representantes da consoante africada surda [tʃ].(12)

O português perdeu, entretanto, o contacto direto com muitas das línguas com que convivera em séculos anteriores e às quais foi buscar muitos termos e topónimos que aportuguesou. Além disso, o português, contrariamente a outras línguas europeias, **não tem** uma tradição de transliterações e transcrições de línguas com outros alfabetos ou sem alfabetos, aparentando estar agora acomodadamente dependente das transliterações e transcrições feitas para outras línguas europeias, como o francês, cada vez mais o inglês e marginalmente o alemão.

Com algumas exceções, os aportuguesamentos são feitos atualmente tendo como **referência** essas transliterações e transcrições(13). Confrontamo-nos, assim, com as seguintes soluções para representar os sons /ʃ/ e /tʃ/, correspondentes aos tradicionais «x» e «ch» portugueses (e galegos):

AFI	DIN 31635	fr	en	de	pt gl
ʃ	š	**ch**	sh	sch	x
tʃ	č	tch	**ch**	tsch	ch

Porém, nos aportuguesamentos baseados nas soluções encontradas para essas línguas nem sempre se faz a **validação** dos sons que estas ortografias estrangeiras procuram representar. Por exemplo, o facto de o «x» não ser utilizado por nenhuma das três habituais línguas de referência faz com que o seu uso possa ser esquecido. Assim, na legislação europeia aparece por vezes a ortografia Cachemira (em vez da tradicional e correta Caxemira) por provável influência da ortografia francesa (Cachemire), onde o «ch» é usado para reproduzir o som /ʃ/, correspondente ao «sh» inglês (Kashmir). A marca em português é usar o «x», letra mal-amada por muitos, talvez pelo facto de a etimologia latina (e grega) lhe poder fazer corresponder também outros sons(14), embora, perante novas palavras, o reflexo seja sempre ler como /ʃ/ um «x» no início ou no meio de uma palavra.

A passagem do som /tʃ/ a /ʃ/, que se verificou na generalidade do vocabulário, verifica-se também nas palavras de origem estrangeira entradas mais recentemente no português, muitas provavelmente já via transliterações e transcrições do inglês, em que o «ch» corresponde a /tʃ/. Em português, a marca dessa (não) pronúncia é dada igualmente pelo «ch». O trígrafo «tch» ainda se pode encontrar na legislação europeia (por exemplo, Tchernobil) por provável influência da transliteração francesa. No Brasil, mais do que em Portugal, há ainda casos de nomes de origem estrangeira em que se força o som /tʃ/ nos aportuguesamentos. É o caso da Tchéquia e por vezes da Tchetchênia ou mesmo do Tchade. Ou seja, no Brasil pode utilizar-se «x» para o som /ʃ/ e «tch» para o som /tʃ/, quando produzido («ch» quando não produzido).

Ver, por exemplo, os aportuguesamentos:

- 1986 Chernobil ru: Чернóбыль; fr: Tchernobyl; en: Chernobyl
- 1991 Chechénia ru: Чечня́; fr: Tchétchénie; en: Chechnya
- 2014 Sóchi ru: Сóчи; fr: Sotchi; en: Sochi

De facto, verifica-se que, mais de um século volvido sobre o Formulário Ortográfico de 1911, as novas palavras já (quase) não nos chegam com ortografia francesa, mas sim com ortografia inglesa. O *Livro de Estilo* de 1997 do *Público* testemunha essa transição, pois ainda recomenda as ortografias «afrancesadas» Tchernobil e Tchetchénia para topónimos que entraram nas notícias ainda no século XX. Quanto ao palco das olimpíadas de inverno de 2014, o *Público* quase já não utiliza a ortografia Sótchi. A tradução portuguesa das instituições europeias na viragem do século também foi testemunha dessa transição da influência francesa para a influência inglesa.

Sempre que houver a necessidade (e a vontade(15)), poderemos **aportuguesar** a grafia das novas palavras (topónimos ou nomes comuns) usando o «x» para a transposição para o português do som /ʃ/ e o «ch» para a transposição para o português do som /tʃ/. Porém, hoje em dia com o recurso a toda a informação multilingue disponível na Internet não estamos limitados às línguas veiculares tradicionais como línguas de partida e **validação** para os aportuguesamentos. Apresenta-se a seguir uma tabela com as soluções ortográficas mais comuns utilizadas em vários alfabetos para representar o som /ʃ/ e o som /tʃ/ em palavras próprias ou transliteradas de outras línguas.

AFI	**DIN 31635**	**fr**	**en**	**de**	**cs**	**ro**	**tr**	**ru tg**	**ar**	**fa**	**hi**	**bn**	**pt gl**
ʃ	š	ch	sh	sch	š	ş	ş	ш	ش	ش	श	শ	x
tʃ	č	tch	ch	tsch	č	ce, ci	ç	ч	تش	چ	च	চ	ch

ISO 639-1: **fr** — francês; **en** — inglês; **de** — alemão; **cs** — checo; **ro** — romeno; **tr** — turco; **ru** — russo; **tg** — tajique; **ar** — árabe; **fa** — farsi; **hi** — hindi; **bn** — bengali; **pt** — português; **gl** — galego

Com base nestes dados, como aportuguesar e grafar em português, por exemplo, os topónimos Душанбе (a capital do Tajiquistão), Чечня́ (uma república da Federação da Rússia) ou نواكشوط (a capital da Mauritânia)? Ou ainda, os estados Chin e Shan (Birmânia)? Estes e outros casos de estudo:

Casos de estudo n.º 1 — línguas indo-iranianas

1.a) tg: Душанбе — Duchambé, Dutchambé ou Duxambé?

O fim da União Soviética, com a consequente independência das repúblicas que a constituíam, como é o caso do Tajiquistão, ocorreu num período em que nas instituições europeias os originais ainda eram predominantemente em francês e em que a principal fonte de toponímia estrangeira era a *Lello Universal* (adaptação portuguesa da Larousse) e em que o acesso à Internet não existia. A grafia portuguesa utilizada até agora nas instituições europeias para a capital do Tajiquistão — Duchambé — decorrerá assim do francês Dou**ch**anbé.

O tajique é uma variante da língua persa (ou farsi), pertencente ao grupo indo-iraniano das línguas indo-europeias. O persa mantém os sons /ʃ/ e /tʃ/. A importância dessa diferença fez com que o alfabeto persa, baseado no alfabeto árabe, inclua o چ, uma letra específica para o fonema /tʃ/, estranho à generalidade das variedades do árabe[16], face a ش (para o som /ʃ/). Devido à incorporação no império russo durante o século XIX, o tajique escreve-se oficialmente com o alfabeto cirílico, embora também se escreva com o alfabeto persa. O cirílico tem letras para os fonemas /ʃ/ (ш) e /tʃ/ (ч).

AFI	**fa** isolado	**fa** inicial	**fa** médio	**fa** final	**tg**	**hi**	**bn**	**pt gl**
ʃ	ش	شـ	ـشـ	ـش	ш	श	শ	x
tʃ	چ	چـ	ـچـ	ـچ	ч	च	চ	ch

Proposta de aportuguesamento: **Duxambé** e não Duchambé, marcando graficamente com «x» o som /ʃ/ do endónimo dado pela letra ш, como o fazem outras línguas de referência!

tg	**fa**	**fr**	**en**	**de**	**es**	**ca**
Душанбе	دوشنبه	Douchanbé	Dushanbe	Duschanbe	Dusambé	Duixanbe

A raiz indo-europeia do tajique pode intuir-se se se souber que Duxambé significa segunda-feira (du «dois» e xambé «dia») no calendário persa[17].

1.b) fa: بلوچستان — Balochistão, Balotchistão ou Baloxistão?

Região de língua persa situada a cavalo entre o Irão e o Paquistão.

A grafia **Balochistão** está correta, marcando o «ch» graficamente o som /tʃ/ do endónimo dado pela letra ـچـ, forma do چ a meio das palavras, como o fazem outras línguas de referência!

fa	**fr**	**en**	**de**	**gl**	**es**	**ca**
بلوچستان	Baloutchistan	Balochistan	Belutschistan	Beluchistán	Baluchistán	Balutxistan

A língua persa deu várias palavras e topónimos ao português com os sons /ʃ/ e /tʃ/:

- xá (fa: شاه; en: *shah*; fr: *chah;* de: *schah;* es: *sah*)
- xaile (fa: شال; en: *shawl*; fr: *châle*)
- xamata — tipo de tecido
- Xerazade (fa: شهرزاد)
- Queixome (fa: قشم; en: Keshm; de: Qeschm) — antiga possessão portuguesa vizinha de Ormuz
- Mexede (fa: مشهد; en: Mashhad; fr: Mechhed) — 2.ª cidade do Irão
- Xiraz (fa: شیراز; en: Shiraz; fr: Chiraz; de: Schiraz) — 5.ª cidade do Irão
- Badaquexão (fa: بدخشان; en: Badakhshan) — região repartida entre o Tajiquistão e o Afeganistão
- chador[18] (fa: چادر; en: *chador*; fr: *tchador*)
- baloche (fa: بلوچ)

1.c) bn: বাংলাদেশ — Bangladeche, Bangladetche ou Bangladexe?

Antigo Paquistão Oriental.

Tal como o farsi, o bengali e o hindi (e o concani de Goa), línguas indo-iranianas, têm os sons /ʃ/ (শ, श) e /tʃ/ (চ, च), refletidos nos seus alfassilabários, o bengali-assamês e o devanágari, respetivamente.

Proposta de aportuguesamento: **Bangladexe** e não Bangladeche, marcando graficamente com «x» o som /ʃ/ do endónimo (শ), como o fazem outras línguas de referência! O vocabulário toponímico do *Vocabulário Ortográfico Comum* (VOC) adota igualmente Bangladexe[19].

bn	hi	en	de	es
বাংলাদেশ	बांग्लादेश	Bangladesh	Bangladesch	Bangladés

Notar o paralelismo com Pradexe (hi: प्रदेश), nome de vários estados da vizinha Índia:

- Andra Pradexe (hi: आन्ध्र प्रदेश) — província dos andras, povo indiano
- Arunachal Pradexe (hi: अरुणाचल प्रदेश) — província das montanhas do levante
- Himachal Pradexe (hi: हिमाचल प्रदेश) — província da neve, cf. Himalaia (lugar da neve)
- Madia Pradexe (hi: मध्य प्रदेश) — província central
- Utar Pradexe (hi: उत्तर प्रदेश) — província do norte

O sânscrito e as atuais línguas indo-iranianas do subcontinente indiano forneceram ao português várias palavras, que refletem essa diferença entre os sons /ʃ/ e /tʃ/. Exemplos:

- Caxemira (hi: कश्मीर; en: Kashmir; fr: Cachemire; de: Kaschmir)
- caxemira (en: *cashemere*; fr: *cachemire*) — tecido de lã
- vaixá ou vaixiá (hi: वैश्य; en: *vaishya*) — casta indiana
- xátria (hi: क्षत्रिय; en: *kshatriya*) — casta indiana
- Xiva (hi: शिव; en: Shiva; de: Schiwa) — divindade hindu
- xivaísmo — ramo do hinduísmo
- achar (hi: अचार; en: *achar*) — picles indianos
- chacra (hi: चक्र; en: *chakra*) — centro de energia do corpo humano
- champó (hi: चंपक ; en: *champak*) — árvore conhecida pela qualidade da madeira
- chapati (hi: चपाती; en: chapati) — tipo de pão
- Chaul (hi: चौल) — antiga possessão portuguesa a sul de Bombaim.

Champô (hi: चॅंपो; cf. en: *shampoo*) é um exemplo interessante, pois o aportuguesamento champô está mais próximo da pronúncia do hindi [tʃãːpoː] do que da pronúncia do inglês [ʃæmˈpuː], que neste caso não reflete a pronúncia original. A variante xampu, mais comum no Brasil é claramente inspirada na grafia inglesa. Terá havido com a grafia champô algum efeito de validação via concani devido à presença portuguesa em Goa, Damão e Diu até 1961?

Casos de estudo n.º 2 — árabe (ou via árabe)

O «x» corresponde ao ش árabe, mas também ao *j* castelhano

- xarope (ar: شراب; gl: xarope; es: *jarabe*)
- xávega (gl: xávega; es: *jábega*)
- enxaqueca (ar: الشقيقة; gl: xaqueca; es: *jaqueca*)
- enxoval (gl: enxoval; es: *ajuar*)
- almoxarife (gl: almoxarife; es: *almojarife*)
- xadrez (ar: شطرنج; gl: xadrez; es: *ajedrez*)
- xária (ar: شريعة; gl: xaria)
- Marraquexe (ar: مراكش; fr: Marrakech; en: Marrakesh; de: Marrakesch; gl: Marraquex)
- Mogadixo (ar: مقديشو; it: Mogadiscio; en: Mogadishu; de: Mogadischu; gl: Mogadixo)
- Daexe (ar: داعش; fr: Daech; en: Daesh; de: Daesch)

2.a) ar: نواكشوط — Nuaquechote, Nuaquetchote ou Nuaquexote?

Capital da Mauritânia, antiga colónia francesa. A origem do topónimo é o berbere Nawākšūṭ (Lugar dos Ventos)[20].

Proposta de aportuguesamento: **Nuaquexote**, por coerência com o /ʃ/ árabe [nwɑːkˈʃɒt], tal como é feito pelo francês (Nouakchott e não Nouaktchott). Confrontar com Chade (ar: تشاد; fr: Tchad), Lago (por referência ao lago Chade), em canúri, língua nilo-sariana.

- **ʃ** ش > x
- **tʃ** تش > ch

2.b) ar: هاشمي — Hachemita, Hatchemita ou Haxemita?

Adjetivo utilizado no nome oficial da Jordânia para designar a casa reinante.

Proposta de aportuguesamento: **Reino Haxemita da Jordânia**, por coerência com o árabe e outras línguas de referência, abandonando o decalque do «ch» francês.

ar	en	fr	de
المملكة الأردنيَّة الهاشميّة	Hashemite Kingdom of Jordan	Royaume hachémite de Jordanie	Haschemitische Königreich Jordanien

2.c) Toponímia de Lisboa e arredores

- Carnaxide (*qarn ax-xidda*)
- Caxias (*kâxih*)
- Pirescoxe (*bîr kûxa*) — poço do forno
- Xabregas (*xabraka*)

Alcabideche (*al-qabidaq*) — mãe de água — não vai contra a regra de aportuguesamento das palavras de origem árabe, pois o «ch» não deriva do som /ʃ/. Há aliás um Alcabideque, em Condeixa-a-Nova, Coimbra (outra zona com muita toponímia de origem árabe). Já Alcochete (*al-kûxât*) — os fornos — não segue a regra. Porém, ainda no século XIX se utilizava a grafia Alcoxete.

Casos de estudo n.º 3 — línguas caucasianas (via russo)

Sendo as línguas caucasianas largamente desconhecidas fora da região, os topónimos do Cáucaso chegam-nos não pelas línguas caucasianas locais, mas sim via russo, língua eslava, onde há diferença clara entre /ʃ/ e /tʃ/.

ru	fr	en	de	gl	es	ca
Чечня́	Tchétchénie	Chechnya	Tschetschenien	Chechenia	Chechenia	Txetxènia
Ингуше́тия	Ingouchie	Ingushetia	Inguschetien		Ingusetia	Ingúixia

3.a) ru: Чечня́ (ce: Нохчийчоь) — Chechénia, Tchetchénia ou Xexénia?

Proposta de aportuguesamento: **Chechénia**, marcando graficamente o som /tʃ/, como o fazem outras línguas de referência.

3. b) ru: Ингуше́тия (inh: ГІалгІайче) — Ingúchia, Ingútchia ou Ingúxia?

Proposta de aportuguesamento: **Ingúxia** (ou **Inguxétia**) e não Ingúchia (ou Inguchétia), marcando graficamente o som /ʃ/, como o fazem outras línguas de referência.

Casos de estudo n.º 4 — línguas tibeto-birmanesas e tai-cadai (via inglês)

Para a Birmânia, sendo as línguas das regiões fronteiriças fora da bacia do Irrauádi[21] largamente desconhecidas fora da região, os termos chegam-nos via transliterações inglesas.

4.a) en: Shan — Chã, Tchã ou Xã?

Proposta de aportuguesamento: **Xã**, usando o inglês como referência e validação.

4.b) en: Chin — Chim, Tchim ou Xim?

Proposta de aportuguesamento: **Chim**, usando o inglês como referência e validação.

4.c) en: Kachin — Cachim, Catchim ou Caxim?

Proposta de aportuguesamento: **Cachim**, usando o inglês como referência e validação.

Casos de estudo n.º 5 — línguas europeias com alfabeto latino

As línguas que usam alfabetos latinos colocam problemas curiosos, pois há a tendência de facilitar nos aportuguesamentos, utilizando-se as mesmas letras, omitindo eventualmente os diacríticos mais ou menos desconhecidos, mesmo com o perigo de, involuntariamente, se chegar a formas foneticamente muito afastadas do endónimo. Um exemplo: em Jugoslávia (sr: Југославија/ hr: Jugoslavija), o «j» servo-croata vale por «i» (facto tido em conta no Brasil — Iugoslávia —, por possível controlo junto da importante comunidade de descendentes de jugoslavos a viver no Brasil).

5.a) ro: Chişinău — Quichinau, Quitchinau ou Quixinau

(de: Kischinau; ru: Кишинёв)

No romeno, a letra «ş» representa o som /ʃ/, como em Chişinău, capital da Moldávia (cf. a grafia em romeno para Duşanbe e Inguşetia). Também tal como em italiano, o som /tʃ/ é representado pela letra «c» antes de «e» ou «i» (exemplo: So**c**i, Sóchi) ou pelo dígrafo «ci» antes das outras vogais (exemplo: **Ci**ad, Chade ou **Cece**nia, Chechénia). Em romeno, tal como em italiano, o dígrafo «ch» corresponde ao som «k».

Proposta de aportuguesamento: **Quixinau** e não Quichinau ou Quitchinau, e muito menos o decalque Chisinau. Foi este o entendimento do vocabulário toponímico do VOC[(22)].

- **ʃ** ş > x
- **tʃ** ci, ce > ch

5.b) cs: Česko — Chéquia, Tchéquia ou Xéquia?

(fr: Tchéquie; en: Czechia; de: Tschechien; es: Chequia; ru: Чéхия)

No checo, e outras línguas eslavas, o som /tʃ/ é representado pela letra «č», como em Česko. Em Portugal escreve-se **Chéquia**, passando o /tʃ/ a /ʃ/, dando o «ch» a indicação dessa pronúncia original (gl: Chequia). Os brasileiros escrevem Tchéquia para forçar a pronúncia do /tʃ/.

- **ʃ** š > x
- **tʃ** č > ch

Em jeito de resumo

Convém haver uma harmonização entre antigos e novos aportuguesamentos. Não se querendo ir para abordagens revolucionárias, como a supressão, mesmo retroativa, da letra «x» e a imposição do «ch» para representar quer o /ʃ/ quer o /tʃ/, a via óbvia parece ser aprender com os termos históricos e adotar aportuguesamentos que tenham em conta as pronúncias dos endónimos (o que era a tradição). O VOC parece tender para esta abordagem.

Muito importante: a utilização do dígrafo «ch» nas transliterações e transcrições francesas e inglesas para representar sons diferentes tem sido fonte de alguma perturbação na criação de novos aportuguesamentos. Mnemónica para futuros aportuguesamentos:

- **ch** francês ou **sh** inglês > **x**
- **ch** inglês ou **tch** francês > **ch**

Paulo.Correia@ec.europa.eu

---

(1) Portal da Língua Portuguesa, *Acordo Ortográfico: Formulário Ortográfico de 1911*, http://www.portaldalinguaportuguesa.org/acordo.php%3Faction%3Dacordo%26version%3D1911.

(2) «**11.** *ch*: Emprega-se como inicial e medial, e nunca como final. Na pronunciação do idioma culto, e bem assim nos vernáculos meridionais, confunde-se no valor há mais de dois séculos com o *x* inicial, do qual se diferença pela origem. Corresponde o *ch*, em geral, a *cl*, *fl*, *pl*, latinos, e a *ch* francês nas palavras desta proveniência; ex.: *chave* (lat. c l a u e m), *chama* (lat. f l a m m a), *chuva* (lat. p l u u i a), *chapéu* (fr. *chapeau*). Corresponde a *ll* e a *ch* castelhanos.

O *ch* com valor de *k* é substituído por *qu* antes de *e*, *i*, e por *c* em qualquer outra situação; ex.: *monarca*, *monarquia*, *querubim*, *côro*, *cloro*, *corografia*, *catecúmeno*, *crisol*. (…)

68. *x*: Esta letra tem cinco valores no idioma comum e literário; são os seguintes:

1.º Como inicial — *xadrez*, *caixa*.

2.º Como *ss* — *auxílio*, *próximo*.

3.º Como *s* — *mixto*, *Félix*.

4.º Como *cs*; *cx* — *fixo*, *sexo*; *córtex*, *sílex*.

5.º Como *(e)is* — *exame*, *êxito*, *texto*.

Nas palavras de origem arábica, e quando é inicial, tem sempre o primeiro valor; ex.: *xabouco*, *axorca*, *xarope*, *elixir*; *Xerxes*, *Xenofonte*, etc.

69. Além desta multiplicidade de valores, alterna, com relação ao primeiro, com o grupo *ch*, o qual, como já se disse, representa *cl*, *fl*, *pl* latinos; assim, temos; *xá* (rei) e *chá* (planta), *xeque* (regedor) e *cheque* (bilhete de banco), *buxo*, lat. b u x u m (planta), e *bucho*, lat. m u s c ' l u m (estômago e músculo).

A consulta ao VOCABULÁRIO é indispensável para o emprêgo de qualquer dêstes dois símbolos, actualmente equivalentes no valor.»

Cf. Imprensa Nacional, *Bases para a unificação da ortografia que Deve Ser Adoptada nas Escolas e Publicações Oficiais*, Lisboa, 1911, http://library.umac.mo/ebooks/b24939092.pdf.

(3) Aportuguesamentos, não transliterações ou transcrições portuguesas, processos para os quais não há tradição consagrada em português. A tradição é adotar a transliteração ou transcrição utilizada pela fonte consultada (inglesa ou, cada vez menos, francesa).
(4) «O alfabeto galego moderno componse das seguintes letras; xunto coa letra e co nome dáse a pronuncia máis representativa:
LETRA (…)**x**(…) NOME (…)xe(…) PRONUNCIA (…)**[ʃ]**, [ks̺] (…)
Na ortografía galega moderna utilízanse tamén os seguintes dígrafos que representan un único fonema:
DÍGRAFO **ch** (…) NOME ce hache (…) PRONUNCIA **[tʃ]** (…)»
Real Academia Galega — *Normas Ortográficas e Morfolóxicas do Idioma Galego*, https://academia.gal/documents/10157/704901/Normas+ortogr%C3%A1ficas+e+morfol%C3%B3xicas+do+idioma+galego.pdf.
(5) Portal da Língua Portuguesa, *Vocabulário Ortográfico Português*, http://www.portaldalinguaportuguesa.org/.
(6) Real Academia Galega, *Dicionario da Real Academia Galega*, https://academia.gal/dicionario/.
(7) Universidade de Santiago de Compostela: Instituto da Língua Galega, *Dicionario de pronuncia da lingua galega*, https://ilg.usc.es/pronuncia/.
(8) Além disso, em muitos casos, os portugueses terão sido os primeiros ocidentais a contactar diretamente esses povos. Por exemplo, quando Camões falou do Barém n'*Os Lusíadas* ainda as grafias Bahreïn (em francês) e Bahrain (em inglês) não haviam sido criadas.
(9) Ver alguns exemplos:

	pronúncia	te	tio	mais	três	dez	casca	cisco
pt	Díli	tɨ	tˈi.ʊ	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kə	sˈiʃ.kʊ
	Lisboa (não padrão)	tɨ	tˈi.u	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.ku
	Lisboa (padrão)	tɨ	tˈi.u	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.ku
	Luanda	tɨ	tˈi.ʊ	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.kʊ
	Maputo (não padrão)	tɨ	tˈi.ʊ	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.kʰʊ
	Maputo (padrão)	tɨ	tˈi.u	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.kʊ
	Rio de Janeiro (não padrão)	**tʃɪ**	**tʃˈi.ʊ**	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.kʊ
	Rio de Janeiro (padrão)	**tʃɪ**	**tʃˈi.ʊ**	mˈajʃ	tɾˈeʃ	dˈɛʃ	kˈaʃ.kɐ	sˈiʃ.kʊ
	São Paulo (não padrão)	ti	tˈi.ʊ	mˈajs	tɾˈes	dˈɛs	kˈas.kə	sˈis.kʊ
	São Paulo (padrão)	**tʃi**	**tʃˈi.ʊ**	mˈajs	tɾˈes	dˈɛs	kˈas.kə	sˈis.kʊ
gl		ˈti	ˈtiọ (tío)	ˈmajs (máis)	ˈtɾes (tres)	ˈdɛθ	ˈkaskạ	ˈθiskọ

O aportuguesamento «quitinete» do inglês *kitchenette*, proposto pelo dicionário Aulete Digital, repousa sobre a pronúncia carioca [ki.tʃɪ nˈe.tʃɪ], não funcionando, por exemplo, para a pronúncia lisboeta [ki.ti.nˈɛ.tɨ].
Cf. Aulete, *quitinete*, http://www.aulete.com.br/quitinete.
A evolução da pronúncia do «s» e «z» para /ʃ/ em finais de frase torna possíveis os aportuguesamentos Tasquente e Bisqueque (em vez de Taxequente e Bixequeque) e terá possibilitado em Goa o aportuguesamento Bardez em vez de Bardexe.
Notar também a solução encontrada pelo mirandês para evitar o /ʃ/ final do português de Portugal:
«D. **Deniç**, an pertués D. Dinis (Lisboua [?], 9 de Outubre 1261 — Santaren. 7 de Janeiro de 1325) fui l sesto rei de Pertual. Filho de D. Fonso III i de la anfanta **Beatriç** de Castielha, nieto de Fonso X de Castielha, fui aclamado an Lhisboua an 1279.»
Biquipédia, *Deniç I de Pertual*, https://mwl.wikipedia.org/wiki/Deni%C3%A7_I_de_Pertual.
(10) «1.9. A GRAFÍA X (…)
A grafía moderna non corresponde á medieval. O galego arcaico tiña un son prepalatal fricativo sonoro, [ʒ], representado por g ou j (g˜esta, geada, gente, janeiro, hoje, beijo), e un son prepalatal fricativo xordo [ʃ], representado por x (coxo, eixe, feixe, dixo, peixe, caixa, queixa, baixar). A diferenza do portugués, do catalán, do francés ou do italiano, que conservan os sons sonoros e teñen que os diferenciar graficamente dos xordos á maneira medieval, o galego eliminou o son sonoro desde fins da Idade Media, polo que se fixo innecesaria a dobre grafía, pois a pronuncia é sempre [ʃ]. Por iso debemos escribir da mesma maneira xente, xaneiro, xesta, hoxe, xeada e coxo, eixe, feixe, dixo, peixe, caixa, queixa etc.»
Real Academia Galega, *Normas Ortográficas e Morfolóxicas do Idioma Galego*, 23.ª ed., março de 2012, https://academia.gal/documents/10157/704901/Normas+ortogr%C3%A1ficas+e+morfol%C3%B3xicas+do+idioma+galego.pdf.
(11) Porto Editora, *Dicionário Infopédia da Língua Portuguesa*, https://www.infopedia.pt/.
(12) Ciberdúvidas da Língua Portuguesa, *Quichinau e Tiráspol*, https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/quichinau-e-tiraspol/32126.
(13) No Brasil, o português, graças aos descendentes da imigração japonesa e sírio-libanesa manteve um contacto mais direto com o japonês e o árabe, o que pode levar a melhores aportuguesamentos no Brasil do que em Portugal para termos com origem nessas línguas, por provável validação por nipofalantes e arabofalantes. Cf. miso [ˈmizu] em Portugal (pela grafia, não pela fonética do inglês) e missô [ˈmiso] no Brasil, mais próximo do japonês (e do inglês).
(14) De notar que no galego esses sons adicionais se limitam ao /ks/, face aos /ks/, /s/, e /z/ portugueses.
(15) Há uma tendência manifestada por alguma imprensa de se parar com qualquer aportuguesamento de topónimos e mesmo com a tradução de alguns termos técnicos ingleses. Segundo essa visão da tradução, o tradutor deverá ser um divulgador da terminologia inglesa junto daqueles que não têm capacidade de ler diretamente a versão inglesa dos documentos. Um dos argumentos é de que a tradução (aportuguesamento) dificulta a pesquisa em inglês dos termos na Internet. Acessoriamente,

essa abordagem isenta o jornalista/tradutor de compreender os conceitos e dificulta a apreensão dos mesmos por boa parte dos leitores.

[16] Wikipedia, *Voiceless postalveolar affricate*, https://en.wikipedia.org/wiki/Voiceless_postalveolar_affricate.

[17] Dias da semana do calendário persa:

farsi	inglês	português	equivalente
شنبه	Shanbeh	xambé	sábado
یکشنبه	Yek-shanbeh	iequexambé	1.ª-feira (domingo)
دوشنبه	Do-shanbeh	duxambé	2.ª-feira
سه‌شنبه	Se-shanbeh	sexambé	3.ª-feira
چهارشنبه	Chahar-shanbeh	chaarxambé	4.ª-feira
پنجشنبه	Panj-shanbeh	panjexambé	5.ª-feira
جمعه	Jom'eh	**juma** (do árabe) ou adiné	congregação (6.ª-feira)

Os nomes dos numerais utilizados nos dias da semana podem dar pistas para a toponímia da região, como por exemplo, panjexambé (5.ª-feira) e Panjabe (panj «cinco, penta» e abe «rio») — Cinco Rios.

[18] Aulete, *chador*, http://www.aulete.com.br/chador.

[19] *Vocabulário Ortográfico Comum da Língua Portuguesa: Vocabulário Toponímico*, «Bangladexe», http://voc.cplp.org/index.php?action=toponyms&act=details&id=TER.142.034.BD.

[20] Curiosamente o topónimo repete-se na Namíbia, onde a capital é Vinduque (Windoek em neerlandês significa também Lugar dos Ventos).

[21] Estados da Birmânia — regiões administrativas habitadas por minorias étnicas:

en	pt	IATE	gentílico	IATE
Chin State	Estado Chim	—	chim (chins)	—
Kachin State	Estado Cachim	3578591	cachim (cachins)	3543078
Kayah State (Karenni State)	Estado Caiá (Estado Caréni)	—	caiá(s) caréni(s)	—
Kayin State (Karen State)	Estado Caim (Estado Carém)	—	caim (cains) carém (caréns)	932702
Mon State	Estado Mom	—	mom (mons)	—
Rakhine State	Estado Arracão	3571141	arracão (arracões)	—
Shan State	Estado Xã	3578593	xã(s)	895570

[22] *Vocabulário Ortográfico Comum da Língua Portuguesa: Vocabulário Toponímico*, «Quixinau», http://voc.cplp.org/index.php?action=toponyms&act=details&id=TER.150.151.MD.CAP.

Toneladas há muitas (parte 2)

Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

O termo **tonelada** corresponde a vários conceitos. A tonelada, para além da óbvia unidade de massa do sistema internacional (1000 quilos), é também uma unidade de massa tradicional, uma unidade de peso (força), uma unidade de volume, uma unidade de energia, etc. Várias destas toneladas já foram analisadas na primeira parte deste artigo[1]. Nesta segunda parte ver-se-á a utilização da tonelada em unidades de conta no domínio da **energia**.

> Há cerca de um século, o consumo energético mundial consistia em dois milhões de toneladas equivalentes de petróleo (tep) diárias. Em contraste, hoje, todos os dias são consumidos 35 milhões tep.[2]

As **unidades de conta** são unidades nas quais se convertem unidades específicas utilizadas para diferentes formas de energia. As unidades de conta tendem a refletir a estrutura do sistema energético. Assim, não é surpreendente que, numa civilização construída em torno do consumo intensivo de energias fósseis, o petróleo (e antes o carvão) seja utilizado como referência para se medir a energia contida em diferentes fontes energéticas. Esse é o caso quando se querem exprimir os totais anuais dos

consumos de energia primária, reduzindo-se todos os combustíveis a uma **quantidade equivalente** (massa equivalente) de um único combustível — o petróleo.

> Unidade de Conta Energética (Unidade Comum)
> Unidade na qual se convertem as unidades específicas utilizadas para as diferentes formas de energia. No sistema SI, a unidade regulamentar é o joule ou o quilowatt-hora; contudo, as unidades fora do sistema SI, unidades de apresentação ditas unidades convencionais, são ainda usadas correntemente; elas são associadas ao emprego de coeficientes de equivalência e permitem adicionar, nos balanços globais, quantidades de energias diferentes; entre as mais correntes encontram-se a tonelada equivalente de carvão (tec) e a tonelada equivalente de petróleo (tep); se bem que não sejam admitidos no sistema SI, a caloria e os seus múltiplos são ainda utilizados, assim como algumas outras unidades físicas fora desse sistema, tal como a British thermal unit (Btu).(3)

Vejam-se os dados apresentados pela Pordata para o consumo de energia primária em Portugal, total e por tipo de fonte de energia, nos primeiros anos da presente década(4). Os valores são arredondados aos milhares de **toneladas equivalentes de petróleo (tep)**.

Anos	**Consumo de energia primária por tipo de fonte (10^3 tep)**						
	Total	**Carvão**	**Petróleo**	**Gás natural**	**Eletricidade (saldo importador)**	**Resíduos industriais (não renováveis)**	**Energias renováveis**
2011	22 109	2 222	10 332	4 483	242	183	4 648
2012	21 482	2 915	9 297	3 950	679	246	4 395
2013	21 461	2 659	9 381	3 769	239	176	5 238
2014	20 921	2 682	9 089	3 486	78	178	5 409
2015	22 060	3 259	9 447	4 097	195	167	4 895
2016	21 684	2 848	9 157	4 340	-437	208	5 568

Fontes: Direção-Geral de Energia e Geologia/Ministério da Economia, Pordata

N.B.: Observa-se, assim, que em Portugal o petróleo garante metade do consumo de energia primária e as renováveis um quarto do consumo, um pouco acima do gás natural. Valores negativos, como é o caso do saldo para a eletricidade em 2016, indicam que nesse ano o país exportou mais do que importou.

A **tonelada equivalente de petróleo (tep)** é a quantidade de energia que se pode obter a partir de uma tonelada de petróleo indiferenciado. Esse valor foi fixado por convenção em 10 gigacalorias (ou em unidades do Sistema Internacional: 41,868 gigajoules ou 11,63 megawatts-hora). Trata-se de um valor convencional, pois diferentes tipos de petróleo bruto apresentam diferentes valores de **poder calorífico**.

O poder calorífico de um combustível é um fator que permite calcular a energia produzida por uma unidade de massa ou volume desse combustível. O valor fixado por convenção para a tonelada equivalente de petróleo corresponde ao do **poder calorífico inferior** (PCI). Para o petróleo considera-se que o poder calorífico superior (PCS) é 5% superior ao PCI. Inclui-se uma breve nota devido aos problemas de tradução dos termos e siglas relacionados com o poder calorífico:

- O **Poder Calorífico Superior (PCS)** traduz o poder calorífico quando o vapor de água formado regressa ao seu estado inicial, isto é, condensa restituindo o calor de vaporização.
- O **Poder Calorífico Inferior (PCI)**, é aquele em que o calor de vaporização não é restituído, ou seja, escapa-se com os gases de combustão pela chaminé, sendo este o que se verifica nas instalações industriais. É este último que caracteriza e viabiliza a aptidão dos materiais/resíduos para a valorização energética.(5)

A documentação da Entidade Reguladora dos Serviços Energéticos (ERSE), da EDP ou da Galp também utiliza esta terminologia(6), evitando termos e siglas resultantes do decalque das variantes em inglês, que se afastam da terminologia há muito consagrada em Portugal:

pt		**en**		**IATE**
poder calorífico	**PC**	*calorific value*	CV	1212662
		heating value	HV	
poder calorífico inferior	**PCI**	*net calorific value*	NCV	1077330
		lower heating value	LHV	
		lower calorific value	LCV	
poder calorífico superior	**PCS**	*gross calorific value*	GCV	1427197
		higher heating value	HHV	
		higher calorific value	HCV	

Para comparar a energia contida em diferentes fontes energéticas é necessário conhecer os **fatores de conversão** em termos de **tep**. Multiplicando uma massa ou volume de uma dessas fontes pelo respetivo fator de conversão obtém-se o correspondente **equivalente de petróleo** expresso em **tep**. Alguns exemplos:

- **antracite**: 0,638 tep/t
  1 t de antracite = 0,638 toneladas equivalentes de petróleo
- **turfa**: 0,186-0,330 tep/t
  10 t de turfa = 1,86-3,30 tep
- **fuelóleo**: 0,984 tep/t
  100 t de fuelóleo = 98,4 tep

Além dos múltiplos e submúltiplos da tonelada equivalente de petróleo, usa-se também uma outra unidade convencional que tem o petróleo como referência: o **barril**. O barril de petróleo tem, por definição, uma capacidade de 42 galões americanos (ou 35 galões imperiais), i.e. 158,9873 litros, de petróleo a 60 ºF (15 ºC) em condições normais. É uma unidade convencional para quantificar reservas de petróleo e produção de jazidas. Convencionou-se que a queima dessa quantidade de petróleo liberta 5,4 GJ (PCI) — o **barril equivalente de petróleo (bep)** —, o que constitui mais uma unidade de comparação de quantidades de diferentes hidrocarbonetos (petróleo e gás natural) ou ainda do ritmo de extração ou refinação dos mesmos — **barris equivalentes de petróleo por dia (bepd)**.

> Angola continua a perder peso na produção de petróleo e gás da Galp Energia, valendo, no final do terceiro trimestre, cerca de 30% do total. De Janeiro a Setembro, a produção total da empresa nos blocos onde está em Angola ascendeu a 10,5 mil barris equivalentes de petróleo por dia (bepd), menos 2,1 mil bepd face ao mesmo período do ano passado.(7)

No domínio da energia usam-se ainda outras unidades convencionais de conta que não têm o petróleo como referência:

- Antes do petróleo, quando a sociedade industrial girava maioritariamente em torno do carvão, a unidade convencional era a **tonelada equivalente de carvão (tec)**. Aceita-se que 1 tec = 7 Gcal.

  > Em valor energético notou-se uma redução nítida nas aquisições, de cerca de 500000 tec (toneladas equivalentes de carvão), provocada fundamentalmente pela boa hidraulicidade do exercício.(8)

- A partir de meados do século XX o conceito de **tonelada equivalente de TNT** ou, simplesmente, **tonelada de TNT**, com os seus múltiplos quilotonelada de TNT ou megatonelada de TNT, foi utilizado para comparar a energia libertada pelas recém-inventadas bombas atómicas com a energia da explosão do TNT. Continua a ser utilizado para descrever explosões de vários tipos. Aceita-se que 1 tTNT = 1 Gcal.

  > Segundo dizem especialistas em segurança de produtos químicos, a primeira explosão foi causada pelo acetileno em chamas e foi equivalente à detonação de três toneladas de TNT. Já a segunda explosão, a grande bola de fogo que foi visível do espaço e se estima ter sido equivalente a 21 toneladas de TNT e produziu um tremor de terra que atingiu 2,9 pontos na escala de Richter, terá sido desencadeada pela reação explosiva do nitrato de amónio à primeira detonação de acetileno.(9)

Algumas unidades de conta no domínio da energia:

	tep	**Gcal**	**MBTU**	**tec**	**t TNT**	**SI**		**IATE**
						GJ	**MWh**	
tonelada equivalente de petróleo (tep)	1	10,000	39,683	1,429	10,00	41,868	11,63	791395
barril equivalente de petróleo (bep)	0,136	1,462	5,799	0,209	1,36	6,118	1,63	1164574
tonelada equivalente de carvão (tec)	0,700	7,000	27,778	1	7,00	29,308	8,14	1407824
tonelada (equivalente) de TNT	0,100	1,000	3,968	0,143	1	4,184	1,16	—

«*Quand la traduction s'en mêle*»…

Se do ponto visto técnico há entendimento sobre o valor de uma **tonelada equivalente de petróleo**, o mesmo não se pode dizer do termo que a representa — sobretudo no plural —, geralmente em resultado da tradução de *tonnes d'équivalent pétrole* (tep) e de *tonnes of oil equivalent* (toe). Exemplos de variantes encontradas no Google, assinalando-se a negrito as variantes igualmente encontradas nas memórias de tradução interinstitucional Euramis e a itálico decalques do inglês:

- **toneladas equivalentes de petróleo**?
- toneladas equivalentes a petróleo?
- toneladas equivalentes em petróleo?
- **toneladas-equivalentes de petróleo**?
- toneladas equivalentes petróleo?
- toneladas de petróleo equivalentes?
- **toneladas de equivalente de petróleo**?
- **toneladas de equivalente a petróleo**?
- toneladas de equivalente em petróleo?
- **toneladas de equivalente petróleo**?
- **toneladas de equivalente-petróleo**?
- **toneladas equivalente de petróleo**?
- toneladas equivalente a petróleo?
- toneladas equivalente em petróleo?
- **toneladas equivalente petróleo**?
- **toneladas equivalente-petróleo**?
- **toneladas de equivalentes de petróleo**?
- toneladas de equivalentes a petróleo?
- **toneladas de equivalentes petróleo**?
- toneladas de equivalentes-petróleo?
- ***toneladas de petróleo-equivalente****?*
- *toneladas de petróleo equivalente?*
- *toneladas de óleo-equivalente?*
- *toneladas de óleo equivalente?*
- *toneladas de óleo equivalentes?*

A questão aqui é o que fazer com «equivalente». Tratá-lo como adjetivo ou como substantivo? A maioria das fontes portuguesas fala tradicionalmente de **tonelada equivalente**, i.e., «equivalente» como adjetivo da unidade. No entanto, também se encontra «equivalente» como substantivo — **equivalente de petróleo** ou **equivalente-petróleo** — e agora também como adjetivo de petróleo — ***petróleo equivalente*** —, o que já não corresponde à sigla consagrada **tep**.

a) O que é uma **tonelada equivalente**? É uma quantidade equivalente de petróleo em termos de energia libertada. É o que se faz sem o explicitar com o TNT — toneladas (equivalentes) de TNT.

> Qual é a quantidade equivalente de petróleo necessária para produzir 7 gigacalorias? A resposta será 0,7 toneladas — 0,7 toneladas equivalentes de petróleo.

b) O que é um **equivalente de petróleo** ou **equivalente-petróleo**? O «equivalente de petróleo» é uma noção menos intuitiva — quantidade de petróleo considerada necessária para produzir uma energia determinada. Mede-se geralmente em toneladas ou seus múltiplos.

> Qual é a quantidade de equivalente de petróleo necessária para produzir 7 gigacalorias? A resposta será 0,7 toneladas — 0,7 toneladas de equivalente de petróleo ou 0,7 toneladas equivalente-petróleo.

c) O que é um **petróleo equivalente**? Aparentemente é um simples decalque do inglês *oil equivalent*. Há quem leve o decalque mais longe e fale mesmo de **óleo equivalente** — toneladas de óleo equivalente (toe).

> Qual é a quantidade de (petr)óleo equivalente necessária para produzir 7 gigacalorias? A resposta será 0,7 toneladas — 0,7 toneladas de (petr)óleo equivalente.

É interessante verificar que o Google mostra uma clara predominância de soluções do tipo «toneladas equivalentes de petróleo». É o caso da generalidade da documentação da ERSE, do Instituto Nacional de Estatística ou da EDP. Já nas memórias de tradução Euramis, em que os conteúdos em português resultam essencialmente da tradução, são igualmente frequentes soluções do tipo «toneladas de equivalente petróleo», pela tradução de «*tonne of oil equivalent*» ou «*tonne d'équivalent pétrole*». O *Código de Redação Interinstitucional* regista: tonelada equivalente-petróleo; tonelada equivalente-carvão. Não é indicado o plural.

Por motivos de coerência terminológica, convirá, tanto quanto possível, harmonizar a designação em português utilizada pelos serviços de tradução das instituições europeia, propondo-se para esse efeito a solução consagrada pelo uso em Portugal, isto é, a **tonelada equivalente de petróleo** (tep).

Mais alguns exemplos de utilização de tep:

> Nos termos do artigo 12.º da Portaria n.º 359/82, de 7 de abril, Regulamento da Gestão do Consumo de Energia, deverá ser adotado o valor de 0,91 tep/10³m³ para o coeficiente de redução a tonelada equivalente de petróleo do gás natural.(10)

> tep (tonelada equivalente de petróleo) e tec (tonelada equivalente de carvão)
> A tep e a tec são unidades de conta. A tep é definida por convenção como sendo igual a 10 000 megacalorias e a tec a 7000 megacalorias. Os seus nomes indicam que são aproximadamente equivalentes à quantidade de calor existente numa tonelada de petróleo e numa tonelada de carvão. Para facilitar cálculos utiliza-se o fator 1,5 tec para 1 tep.(11)

> tonelada equivalente de petróleo, Unidade de medida de consumo de energia: 1 TEP = 0,041868 TJ. Os fatores de conversão (tep/tonelada) adotados pela Agência Internacional de Energia (AIE), para 1991, são os seguintes: Gasolina para motor 1,070; Gasóleo/*diesel* 1,035; Fuelóleo pesado 0,960; Gás de petróleo liquefeito 1,130; Gás natural 0,917; o fator de conversão utilizado pela AIE para a eletricidade é: 1 TWh = 0,086 Mtep.(12)

> Tonelada equivalente de petróleo (tep): Unidade de energia. A tep é utilizada na comparação do poder calorífero de diferentes formas de energia com o petróleo. Uma tep corresponde à energia que se pode obter a partir de uma tonelada de petróleo padrão.(13)

Por vezes é necessário traduzir a expressão *oil equivalent* ou *oil equivalents* sem indicação de unidades. Em coerência com o termo tonelada equivalente de petróleo, tal consegue-se utilizando expressões do género **quantidades equivalentes de petróleo**. Em alternativa, poderá ser a ocasião perfeita para usar o termo **equivalente de petróleo** (ou equivalente-petróleo). Exemplos:

en[14]	pt	fr
The average daily net imports to be taken into account shall be calculated on the basis of the crude oil equivalent of imports during the previous calendar year, determined in accordance with the method and procedures set out in Annex I.	As importações líquidas diárias médias a ter em conta são calculadas com base no equivalente de petróleo bruto das importações no ano civil anterior, estabelecido segundo o método e os procedimentos enunciados no anexo I.	*Les importations journalières moyennes nettes à prendre en compte sont calculées sur la base de l'équivalent en pétrole brut des importations durant l'année civile précédente, établie selon la méthode et les modalités exposées à l'annexe I.*

Proved reserves replacement ratio is the extent to which the year's production has been replaced by proved reserves added to our reserve base. The ratio is expressed in oil-equivalent terms and includes changes resulting from discoveries, improved recovery and extensions and revisions to previous estimates, but excludes changes resulting from acquisitions and disposals.[15]

… a taxa é expressa em quantidade equivalente de petróleo
… a taxa é expressa em equivalente de petróleo

UKCS petroleum reserves and discovered resources are 70% oil and 30% gas, when expressed in oil equivalent terms.[16]

… quando expressos em quantidade equivalente de petróleo
… quando expressos em equivalente de petróleo

In all cases, this shows cumulative production, in oil equivalent terms, divided by the currently stated hydrocarbons-initially-in-place, again with all phases and hydrocarbon types combined.[17]

… produção cumulativa em quantidade equivalente de petróleo
… produção cumulativa em equivalente de petróleo

en	pt	fr	IATE
oil equivalent	equivalente de petróleo (equivalente-petróleo)	*équivalent pétrole*	1408593
coal equivalent	equivalente de carvão (equivalente-carvão)	*équivalent charbon*	791263
TNT equivalent	equivalente de TNT (equivalente-TNT)	*équivalent TNT*	1085671

Na parte 3…

… deste artigo veremos como designar em português uma unidade de conta para comparação do potencial de aquecimento planetário de diferentes gases com efeito de estufa.

(continua…)

Paulo.Correia@ec.europa.eu

---

(1) «Toneladas há muitas (parte 1)», *in* «a folha», n.º 58 — outono de 2018, http://ec.europa.eu/translation/portuguese/magazine/documents/folha58_pt.pdf.
(2) Galp Energia, *Em busca de mais e melhor energia — Relatório de Sustentabilidade 2013*, http://www.galpenergia.com/PT/investidor/Relatorios-e-resultados/relatorios-anuais/Documents/Relatorio-de-sustentabilidade_2013.pdf.
(3) EDP, *Glossário*, «T», https://www.edp.com/pt-pt/glossario?page=43.
(4) Pordata, *Ambiente, Energia e Território*, «Consumo de energia primária: total e por tipo de fonte de energia», https://www.pordata.pt/Portugal/Consumo+de+energia+prim%C3%A1ria+total+e+por+tipo+de+fonte+de+energia-1130.
(5) Caracol, P. M. de O., *Avaliação da viabilidade dos combustíveis derivados de resíduos: Caso de estudo da indústria cimenteira*, https://fenix.tecnico.ulisboa.pt/cursos/mec/dissertacao/565303595501234.
(6) Exemplos:

- **Poder Calorífico Inferior (PCI)**
  O poder calorífico inferior é a quantidade de calor liberta pela combustão completa de uma unidade de combustível, admitindo-se que o vapor de água não se encontra condensado.

Entidade Reguladora dos Serviços Energéticos (ERSE), *Glossário*,
http://www.erse.pt/pt/glossario/Paginas/glossario.aspx.

- **Poder Calorífico Superior (PCS)**
  Quantidade de calor produzida na combustão completa, a pressão constante, de uma unidade de massa ou de volume do gás combustível, considerando que os produtos de combustão cedem o seu calor até atingirem a temperatura inicial dos reagentes e que toda a água formada na combustão atinge o estado líquido.
  Entidade Reguladora dos Serviços Energéticos (ERSE), *Glossário*,
  http://www.erse.pt/pt/glossario/Paginas/glossario.aspx.
- **Poder Calorífico Inferior (PCI)**
  O poder calorífico inferior é a quantidade de calor liberta pela combustão completa de uma unidade de combustível, admitindo-se que o vapor de água não se encontra condensado.
  EDP, *Glossário*, «Poder Calorífico Inferior (PCI)», https://www.edp.com/pt-pt/glossario?page=29.
- **Poder Calorífico Superior (PCS)**
  O poder calorífico superior é a quantidade de calor liberta pela combustão completa de uma unidade de combustível encontrando-se o vapor de água condensado e o calor recuperado.
  EDP, *Glossário*, «Poder Calorífico Superior (PCS)», https://www.edp.com/pt-pt/glossario?page=29.
- Dados de referência do Gás Natural

	[Magrebe] (% mol.)	**[GNL] (mín.)**	**Média (máx.)**
	kWh/m³	kWh/m³	kWh/m³
P.C.I. (Poder Calorífico Superior)	11,8	12	11,9
P.C.S. (Poder Calorífico Inferior)	10,66	10,85	10,755

GALP Distribuição Gás Natural, *O que é o Gás Natural?*, https://galpgasnaturaldistribuicao.pt/gas-natural/o-que-e.

(7) Rodrigues, D., «Angola perde peso na produção de petróleo e gás da Galp», *Expansão*, 4.11.2014,
http://expansao.co.ao/artigo/51845/angola-perde-peso-na-producao-de-petroleo-e-gas-da-galp-?seccao=exp_tec.

(8) EDP, *Exercício 88: Documentos de Prestação de Contas do Exercício de 1988*,
http://www.colecoesfundacaoedp.edp.pt/Nyron/Library/Catalog/winlibimg.aspx?skey=A5CB6F7672C542D3A3962946EA86FD45&doc=3031&img=181217.

(9) Ribeiro, F., «Explosões em Tianjin podem ter sido causadas pela água dos bombeiros», *Público*, 14.8.2015,
https://www.publico.pt/2015/08/14/mundo/noticia/explosoes-em-tianjin-podem-ter-sido-causadas-pela-agua-dos-bombeiros-1705023.

(10) Despacho n.º 3157/2002 (2.ª série), https://dre.pt/application/conteudo/2136653.

(11) EDP, *tep (tonelada equivalente de petróleo) e tec (tonelada equivalente de carvão)*,
https://www.edp.com/pt-pt/tep-tonelada-equivalente-de-petroleo-e-tec-tonelada-equivalente-de-carvao.

(12) Instituto Nacional de Estatística, *Sistema de Metainformação*, «Conceito: 1055 – Tonelada Equivalente de Petróleo (TEP)», http://smi.ine.pt/Conceito/Detalhes/1442.

(13) Agência Nacional de Energia Elétrica, *Atlas de Energia Elétrica do Brasil: Fatores de Conversão*, 3.ª ed, 2008,
http://www.aneel.gov.br/documents/656835/14876406/2008_AtlasEnergiaEletricaBrasil3ed/297ceb2e-16b7-514d-5f19-16cef60679fb.

(14) Diretiva 2009/119/CE do Conselho, de 14 de setembro de 2009, que obriga os Estados-Membros a manterem um nível mínimo de reservas de petróleo bruto e/ou de produtos petrolíferos,
https://eur-lex.europa.eu/legal-content/EN-PT-FR/TXT/?uri=CELEX:32009L0119&from=PT.

(15) BP, *Annual Report and Form 20-F 2014*,
https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/investors/bp-annual-report-and-form-20f-2014.pdf.

(16) «OGA: UK oil and gas reserves enough for 20+ years of production», *Offshore Energy Today*, 8.11.2018,
https://www.offshoreenergytoday.com/oga-uk-oil-and-gas-reserves-enough-for-20-years-of-production/.

(17) Stronach, N., «Benchmarking Oil and Gas Field Recovery Targets and Investment in Norway»,
http://gaffney-cline-focus.com/benchmarking-oil-and-gas-field-recovery-targets-and-investment-in-norway.

A influência árabe na língua portuguesa

Anabela Pereira
Direção Geral da Tradução — Comissão Europeia

[Versão portuguesa de um texto francês publicado na Intranet da DGT no âmbito da comemoração do Ano Europeu do Património Cultural.]

Uma língua é como um ser vivo em constante evolução e cada uma reflete uma herança que remonta a tempos imemoriais, fruto de inúmeras influências ao longo do tempo, resultantes do inevitável cruzamento de populações de origens díspares. O português não constitui uma exceção e veicula informações preciosas que refletem os diferentes contributos que se foram sucedendo ao longo da sua história(1).

Os diferentes substratos da língua portuguesa

Pela sua fonética, morfologia, sintaxe e léxico, a língua portuguesa é essencialmente o produto de uma evolução orgânica do latim vulgar, introduzido na península Ibérica no século III a.C. pelos colonos romanos(2). Mas incluía, à partida, um substrato céltico-lusitano correspondente às línguas faladas pelos povos pré-romanos da parte ocidental da península. Após a queda do Império Romano e as invasões bárbaras, o português arcaico desenvolveu-se como um dialeto românico denominado galaico-português. A chegada dos árabes à península trouxe à língua portuguesa uma nova componente(2).

O árabe:* lingua franca *durante séculos

Os árabes desembarcaram na península Ibérica há mais de mil e trezentos anos, em 711, na sequência de uma expansão migratória com origem no Norte de África. A sua civilização, culturalmente mais sofisticada do que a dos povos romano-visigodos presentes na península, exerceu uma influência importante em diferentes domínios e, em particular, sobre a língua. Apesar de o uso do árabe nunca ter sido imposto, esta língua semítica vai desempenhar durante muito tempo o papel de *lingua franca* e ser utilizada como língua veicular durante cinco séculos na ciência, na cultura, na administração e no comércio(1).

A origem de Portugal como nação é uma consequência do movimento de reconquista cristã, lançado no decurso do século XII. Todavia, o prestígio da cultura árabe continuará bem visível na língua, na toponímia, na arquitetura e nas ciências, mesmo após o termo do reino árabe.

O papel de preservação do saber da tradução árabe dos antigos autores

Pouco tempo depois da conquista das regiões que tinham pertencido ao Império Bizantino, os árabes deram início à tradução sistemática dos tratados científicos e filosóficos de autores gregos e romanos. Para os letrados da época cristã posterior à reconquista, essencialmente clérigos e monges, tornou-se bem clara a importância deste contributo dos árabes para a preservação das versões originais de obras antigas. Muitas das versões originais traduzidas pelos árabes tinham efetivamente desaparecido ao longo dos séculos(1).

O desenvolvimento intelectual árabe na origem de numerosos neologismos

Os árabes não se limitaram a compilar e a traduzir todo este património; através de uma intensa atividade científica, filosófica e cultural, contribuíram também consideravelmente para o seu enriquecimento. Este desenvolvimento intelectual conduziu à expansão dos domínios do conhecimento e à emergência de novos conceitos, que deram origem à necessidade de criar novos vocábulos(1).

A língua árabe caracteriza-se por uma grande plasticidade que permite a criação de novos vocábulos com base em radicais que a estruturam. Deste modo, foram criados em árabe milhares de novos termos derivados que não tinham correspondência nem nas línguas clássicas nem nas línguas românicas faladas nessa época[(1)].

A contrafação das primeiras traduções cristãs

Aos primeiros tradutores cristãos colocou-se a questão de saber como traduzir para o grego ou o latim estes novos termos árabes. A opção seguida por estes tradutores foi de fabricar novos vocábulos «romanizados» a partir da raiz semítica, omitindo qualquer explicação etimológica. A intenção clara desta contrafação linguística era dissimular o facto de estes termos terem uma origem árabe, a língua do povo vencido![(1)]

Os investigadores contemporâneos tentaram fazer justiça a esta língua de partida estudando a origem etimológica de uma parte do vocabulário português atual. Este estudo foi nomeadamente baseado nos manuscritos de um autor acima de qualquer suspeita: o franciscano Diego de Guadix (século XVI). Confessor da Inquisição espanhola e originário da região de Granada, onde a língua e as tradições árabes se encontravam ainda bem enraizadas, Diego de Guadix detinha um conhecimento profundo do árabe que era praticamente a sua língua materna. No seu *Diccionario de arabismos* coloca em evidência a influência deste idioma no desenvolvimento da língua latina, que tinha assimilado numerosas raízes semíticas[(1)].

Uma influência que resistiu ao desgaste do tempo

Numerosos vocábulos ainda hoje utilizados no quotidiano testemunham claramente o impacto do árabe sobre a língua portuguesa. Com uma forte presença na toponímia do país, e sobretudo no centro e sul do país, os vocábulos de origem árabe encontram-se igualmente em diversos domínios do conhecimento. A título de exemplo, o termo «Algarve» significa simplesmente «ocidente» em árabe!

O contributo árabe diz essencialmente respeito à esfera material: produziu sobretudo substantivos e, em menor grau, adjetivos. Todavia, os advérbios são muito raros e os verbos quase inexistentes. São igualmente praticamente inexistentes as expressões relativas a qualidades morais e outras noções abstratas.

A influência do árabe exerceu-se ao nível do léxico mas não da estrutura, que se manteve latina[(3)].

No português contemporâneo, numerosos vocábulos de origem árabe são facilmente identificáveis pela presença do artigo definido invariável em árabe ***al***[(4)]. Como os árabes constituíram a classe dominadora e militar, encontram-se numerosos termos relativos a instituições jurídicas e sociais, bem como à arte da guerra. Encontram-se igualmente termos relativos à cozinha e aos alimentos, às indústrias e ao comércio, à agricultura, às ciências e às técnicas, às artes, aos ofícios, ao vestuário, aos animais e plantas, topónimos (sobretudo no centro e sul do país), etc., quase todos anteriores ao século XIII.

É possível classificar as palavras portuguesas de origem árabe de acordo com as seguintes categorias semânticas[(5)]:

- designação de cargos e dignidades: alcaide, alferes, almoxarife;
- termos castrenses: arraial, arrebate, alcácer, alcáçova, atalaia;
- termos administrativos: aldeia, arrabalde, alfoz; alfândega, alvará, almoeda;
- designação de particularidades topográficas: albufeira, alverca, algar, lezíria, recife;
- termos de arquitetura: aljube, chafariz, açoteia, alvenaria;
- termos de ciências exatas: algarismo, álgebra, cifra, auge, etc.
- designação de unidades de medida: almude, arrátel, alqueire, arroba;

- designação de profissões e indústrias: alfaiate, alveitar, almocreve, alvanel, algoz, azenha, atafona, adobe;
- designação de produtos agrícolas e industriais: azeite, álcool, alcatrão;
- termos da vida pastoral: zagal, alfeire, rês, tabefe, almece;
- designação de plantas cultivadas e silvestres: arroz, algodão, alcachofra, cenoura, laranja, açúcar, alfarroba, alecrim, açucena, alfazema;
- designação de animais: atum, alcatraz, alforreca, alacrau, javali;
- designação de artigos de luxo e instrumentos de música: almofada, alcatifa, marfim, alfinete, adufe, rabeca, anafil, alaúde.

Uma curiosidade da língua portuguesa é a interjeição ***oxalá***, que deriva do étimo إن شاء الله (*law xâ allâh*) — se Deus quiser! — e que é utilizado com um sentido idêntico em ambas as línguas.

A tolerância linguística do português às influências externas

O português atual é o fruto de um trabalho multissecular de elaboração e de seleção, de uma evolução específica diversa da das outras línguas românicas. É o reflexo de circunstâncias históricas, de intercâmbios, de mudanças sociais e políticas que forjaram a nação portuguesa e lhe conferiram a sua identidade.

Após ter adquirido o estatuto de língua nacional, o português continuou a evoluir através de um processo de osmose com outras culturas e outras línguas. As descobertas marítimas colocaram os portugueses em contacto com outros falares exóticos, nos quatro cantos do mundo. Estes múltiplos contactos abriram o português a uma tolerância linguística que conduziu à assimilação de inúmeros vocábulos que se encontram ainda hoje presentes no português contemporâneo(1).

Anabela.Pereira@ec.europa.eu

---

(1) Alves, A., *Dicionário de Arabismos da Língua Portuguesa*, Imprensa Nacional-Casa da Moeda e A. Alves, 2013.
(2) Silva, D., «O Português e Suas Influências Linguísticas», *Estudo Prático*, 3.7.2017 (atualização 14.11.2018), https://www.estudopratico.com.br/o-portugues-e-suas-influencias-linguisticas/.
(3) Houaiss, A., «As Projeções da Língua Árabe na Língua Portuguesa: Conferência para o Centro de Estudos Árabes da USP», 1986, Adum, C. N. (transcrição), *Editora Mandruvá e Revistas do CEMOrOc*, http://www.hottopos.com/collat7/houaiss.htm.
(4) Fonseca, F. V. P. da, «Influência árabe na língua portuguesa», *Ciberdúvidas da Língua Portuguesa*, 15.3.2000, https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/influencia-arabe-na-lingua-portuguesa/5010.
(5) Lessa, L. G., «O legado árabe à língua portuguesa», *A Gazeta do Acre*, 23.1.2013, https://agazetadoacre.com/o-legado-arabe-a-lingua-portuguesa/.

Redação: Paulo Correia (Comissão); Valdemar Ferreira (PE); Manuel Leal (Conselho da UE); Victor Macedo (CESE-CR); António Raúl Reis (Serviço das Publicações)
Grupo de apoio: Ana Luísa Faria (Conselho da UE); Hilário Leal Fontes (Comissão); Luís Seabra (Comissão); Susana Gonçalves (Comissão); Ana Lorenzo Garrido (Comissão); Joana Seixas (CESE-CR)
Paginação: Susana Gonçalves (Comissão)
Envio de correspondência: dgt-folha@ec.europa.eu

Edição impressa: oficinas gráficas do Serviço de Infraestruturas e Logística — Bruxelas (Comissão)
Edição eletrónica: sítio Web da Direção-Geral da Tradução da Comissão Europeia no portal da União Europeia — http://ec.europa.eu/translation/portuguese/magazine

Centro de Informação Europeia Jacques Delors
DGAE/Ministério dos Negócios Estrangeiros

A coleção completa d'«a folha» está disponível no catálogo em linha da Biblioteca Jacques Delors
https://infoeuropa.eurocid.pt/pesquisar/wti=(a+folha)+AND+(wfmt=se+OR+wfmt=an)/catalogo=bibliografico

«a folha» ISSN 1830-7809

ISSN 1830-780-9
9 771830 780004